Influence of Gauss and Gauss-Lobatto quadrature rules on the accuracy of a quadrilateral finite element method in the time domain.

Marc Duruflé, Pascal Grob, Patrick Joly

Influence of Gauss and Gauss-Lobatto quadrature rules on the accuracy of a quadrilateral finite element method in the time domain.

M. Durufle, P. Grob, P. Joly
POEMS group. INRIA-Domaine Voluceau
Rocquencourt, France

In this paper, we examine the influence of numerical integration on finite element methods using quadrilateral or hexahedral meshes in the time domain. We pay special attention to the use of Gauss-Lobatto points to perform mass lumping for any element order. We provide some theoretical results through several error estimates that are completed by various numerical experiments.

Keywords: Numerical integration, error estimates, mass lumping, quadrilateral elements

INTRODUCTION

In order to discretize an evolution problem described with partial differential equations, a classical space-time method consists of, whenever possible, using the finite element method for the space approximation and finite differences for the time approximation. Finite element methods involve a mass matrix which has to be inverted at each time step (even with an explicit time scheme). For large scale numerical simulations, this is computationally expensive. The mass lumping technique consists of replacing the *exact non-diagonal* mass matrix by an *approximate diagonal* one [14],[7], which leads to a truly explicit scheme after time discretization. This operation can be performed by using an appropriate quadrature formula (element by element) for the calculation of the integrals defining the entries of the mass matrix. In the case of standard Lagrange finite elements, this simply requires to use the degrees of freedom (i.e. the nodal interpolation points) as quadrature points. Of course, this additional approximation should be done without affecting the accuracy of the method or at least the order of approximation. This imposes some constraints to be satisfied by the quadrature formula (see for instance [4]). In [6], the authors dealt with the case of the 1D transient wave equation and showed that the use of Gauss-Lobatto points as degrees of freedom provides a mass lumping at any order of approximation: this permits us to use Gauss-Lobatto formulas and leads to spectral finite elements. In 2D, the problem is more complicated for P_k Lagrange elements on

Numerical Methods for Partial Differential Equations ???, 1–27 (???)
 CCC ???

triangular meshes since the equivalent of Gauss-Lobatto formulas is not available. It appears that the use of a quadrature formula must be combined with an enrichment of the local finite element space. In [15], the authors constructed triangular finite elements compatible with mass lumping for orders 2 and 3 and more recently in [17], this was extended up to order 6 in 2D and 4 in 3D. However, the techniques employed do not seem to be easily generalized to higher orders. To our knowledge, the construction of a triangular finite element which enables mass lumping for any order is still an open question.

However, since a quadrilateral (or hexahedral in 3D) finite element is constructed by 'tensorization' of a 1D finite element, tensorized Gauss-Lobatto points combined with Lagrange Q_k isoparametric elements, achieve mass lumping for any order, at least for uniform meshes [7]. This idea has been developped and implemented for many physical applications, especially for second order hyperbolic models, i. e. of the form

$$\frac{\partial^2 u}{\partial t^2} + \mathcal{A}u = 0 \tag{0.1}$$

where $\mathcal{A}$ denotes some positive symmetric second order differential operator, used for the simulation of linear wave propagation phenomena. Examples of its application include the 2D acoustics wave equation [15] , the 3D linear elasticity system [5], the Maxwell equations [18], a poro-elasticity model [2] and more recently the Reissner-Mindlin plate model [13]. These papers essentially highlight the interest of the method in terms of computation time and accuracy through numerical tests and numerical dispersion analyses. They emphasize that a reformulation of the standard finite element formulation of the second order problem as a particular first order mixed formulation leads to an implementation that induces a huge gain in terms of memory storage and computational time. In particular, a comprehensive work including numerical tests of convergence is conducted in [10].

Beyond the question of the mass matrix, the computation of the stiffness matrix, induced by the operator $\mathcal{A}$ requires also, in practice, the use of a quadrature formula (that no longer needs to coincide with the one used for the mass matrix). This is even required with the use of general quadrilateral or hexahedral matrices (isoparametric elements): in such a situation, the basis functions of the finite element space are no longer locally polynomial and the entries of the stiffness matrix can not be determined analytically, even in the case of constant coefficients. The question of the preservation of the order of approximation is also raised by the use of this second quadrature formula. A priori, the most natural choice is to use the Gauss formulas (exact in Q_{2r+1}). However we have also investigated the use of the same Gauss-Lobatto formulas for the mass matrix, as done in [5], even if it can deteriorate the order of accuracy. The interest is that the resulting approximate stiffness matrix leads to more efficient algorithms for the matrix-vector product.

The goal of this paper is to contribute the theoretical justification for the cited works, by analyzing the error due to the approximation of mass and stiffness matrices. Such an error analysis is well known in the case of triangular or rectangular elements for which the transformation between the reference element and where the current element is an affine map [4]. The analysis is much less obvious in the case of isoparametric elements for which functions of the approximation space are now rational functions in each ele-

ment, instead of polynomials. This makes the analysis of the interpolation error more tricky [1],[11] and the analysis of the quadrature error more complex. Some results in this direction are given in [3] for some isoparametric elements provided that complicated assumptions (which are difficult to check) are satisfied by the quadrature rule and the approximation space.

To summarize our main results, a table is provided in figure 1, which gives the convergence rate in the energy norm, depending on the space dimension, and on the quadrature rule used for the stiffness matrix. This table leads to the following comments :

	Gauss	Gauss-Lobatto
d = 2	h^r	h^{r-1}
d = 3	h^{r-1}	h^{r-2}

FIG. 1. Theoretical rates of convergence according to space dimension and quadrature rule for the stiffness matrix. The mass matrix is computed with Gauss-Lobatto formulas in order to obtain mass-lumping

- The numerical experiments suggest that our estimates for Gauss-Lobatto are not optimal : one conjectures an extra power of h.
- The loss of one power of h when passing from 2-D to 3-D, confirmed by the numerical results, is due to the approximation of the mass matrix.
- In the case of non-distorted meshes (parallelograms, parallepipeds), the $O(h^r)$ optimal accuracy is proved for any dimension and quadrature rule.
- The use of Gauss-Lobatto rules to compute the stiffness matrix induces a computation time twice smaller than with Gauss rules.

Our analysis is based on an approach that is very close to the one used in [15] for triangular elements. It consists of establishing error estimates in the time harmonic domain for a family of elliptic problems using Strang's Lemma [4] and exploiting these results in the time domain using the inverse Laplace transform. The outline of our paper is as follows:

1. In section I, we present our finite element scheme on a model problem, namely the scalar wave equation, and introduce some tools and notations needed for the analysis.

2. Section II is devoted to the error analysis. As mentioned above, we start by studying in subsection A a family of elliptic problems in the time harmonic domain. We address the question of the influence of numerical integration, considering separately, for technical reasons that will made be clear later, the case of the mass matrix and the case of the stiffness matrix. In subsection B, we establish the time domain error estimates.

3. In section III, we present a thorough numerical study in order to illustrate our theoretical results.

I. PRESENTATION OF THE FINITE ELEMENT METHOD

We introduce the following model problem :

Find $u : \Omega \times [0,T) \longrightarrow \mathbf{R}$ solution of

$$\begin{cases} \dfrac{\partial^2 u}{\partial t^2}(x,t) - \Delta u(x,t) &= f(x,t) \text{ in } \Omega \times [0,T) \ , \\ \dfrac{\partial u}{\partial n}(x,t) &= 0 \qquad \text{in } \partial\Omega \times [0,T) \ , \\ u(x,0) = \dfrac{\partial u}{\partial t}(x,0) &= 0 \qquad \text{in } \Omega \ . \end{cases} \tag{1.1}$$

and its variational formulation (for $f \in L^2(0,T;L^2(\Omega))$):

Find $u \in C^1(0,T;L^2(\Omega)) \cap C^0(0,T;H^1(\Omega))$ solution of

$$\frac{d}{dt^2}\int_\Omega u(t)\, v + \int_\Omega \nabla u(t) \cdot \nabla v = \int_\Omega f(t)\, v \ , \ \forall\, v \in H^1(\Omega) \ . \tag{1.2}$$

In this section, we present the finite element scheme based on spectral elements and accurate quadrature formula that allow for mass lumping. As mentioned in the introduction, the aim of the paper is the analysis of this scheme. To this end, we need to define precisely some important aspects and notations related to the geometry of the mesh, the choice of the finite element (degrees of freedom, finite element space, etc), and the quadrature rules used in the next sections.

A. The isoparametric finite element $\mathcal{Q}^r(\widehat{K}) = (\widehat{K}, \mathbb{Q}_r(\widehat{K}), \widehat{\Xi}^d)$ in quadrilaterals and hexahedra

The first difficulty is to adopt a good mathematical definition of a quadrilateral in 2D or an hexahedron in 3D. This is less clear than it seems at a first glance, especially in 3D, a domain where the finite element literature is rather poor or imprecise.

For application to finite element, it is useful to define a quadrilateral or an hexahedron as the image by a multilinear map of the unit square (in 2-D) or cube (in 3-D), namely

$$\widehat{K} = [0,1]^d \tag{1.3}$$

Let $\mathbb{Q}_r(\hat{K})$ be the set of polynomials of degree r in each variable.

$$\mathbb{Q}_r = span[x_1^{\alpha_1} \cdots x_d^{\alpha_d}], \quad \alpha \in \mathcal{S}_1 = \{\alpha \in \mathbb{N}^d, \quad max\ \alpha_\ell \leq r\}. \tag{1.4}$$

We shall define the admissible $\mathbb{Q}_1$ transforms from $\mathbf{R}$ into itself as,

$$\mathbb{Q}^d_{1,ad} = \{F \in \mathbb{Q}^d_1 \ / \ F|_{\widehat{K}} \text{ is injective } \}, \tag{1.5}$$

We define an element K of $\mathcal{Q}$ as the image by an admissible $\mathbb{Q}_1$ transformation of $\widehat{K}$.

$$K \in \mathcal{Q} \Longleftrightarrow \exists\ F_K \in \mathbb{Q}^d_{1,ad} \ / \ K = F_K(\widehat{K}) \tag{1.6}$$

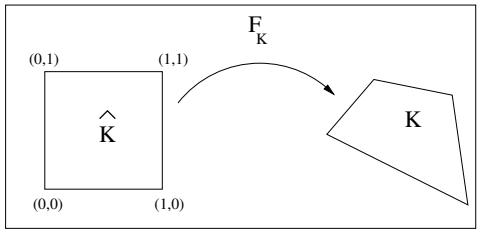


FIG. 2. F_K for two dimensions.

An illustration of the transformation F_K can be seen in the figure 2. By definition, the vertices of K are the images by F_K of the vertices of $\widehat{K}$. In other words, the points x_α are defined by:

$$x_\alpha = F_K(\widehat{x}_\alpha), \quad \text{where } \widehat{x}_\alpha = (\alpha_1, \cdots, \alpha_d) \tag{1.7}$$

Similarly, an edge of K is the image by F_K of an edge of $\widehat{K}$ and, in 3D, a face of K is the image by F_K of a face of $\widehat{K}$. Clearly, the boundary of K is made of the union of its edges in 2D and of its faces in 3D. In 2D, any $K \in \mathcal{Q}$ is a quadrilateral in the usual sense, i.e., with straight edges. In 3D, if each edge of K is a segment, a face of K is in general not planar but curved. As a consequence K is not necessarily convex (while it is in two dimensions). In that sense, K is not necessarily a hexahedron in the usual sense (i.e., with plane faces): our definition generalizes the classical one.

We shall denote F_K^i the coordinates of F_K and by x_α^i the coordinates of $\widehat{x}_\alpha$. An important observation is the following (it can be checked "by hand"):

$$\begin{cases} \forall\ (i,j) \in \{1, \cdots, d\}^2, \quad \dfrac{\partial F_K^i}{\partial \widehat{x}_j}(\widehat{x}) \text{ is a polynomial in } \mathbb{Q}_1 \text{ whose coefficients} \\ \\ \text{are of the form } x_\alpha^\ell - x_\beta^\ell \quad \text{with } \ell \in \{1, \cdots, d\}, \quad (\alpha, \beta) \in S_1^2, \quad \alpha \neq \beta. \end{cases} \tag{1.8}$$

As a definition of the "size" h_K of K, we shall use the maximal distance between vertices:

$$h_K = \sup\ \{|x_\alpha - x_\beta|,\ (\alpha, \beta) \in S_1^2\ \} \tag{1.9}$$

Because of the possible non-convexity of K in 3D, h_K does not necessarily coincide with $diam(K) = \sup\{\ |x-y|,\ (x,y) \in K^2\}$.

If we denote by DF_K the Jacobian matrix of F_K, we set

$$\|DF_K\|_\infty = \sup_{\widehat{x} \in \widehat{K}}\ |DF_K(\widehat{x})|, \quad |DF_K(\widehat{x})| = \sup_{u \neq 0} \frac{|DF_K(\widehat{x})\ u|}{|u|} \tag{1.10}$$

It immediately follows from (1.8) that there exists a constant $C(d) > 0$ such that:

$$\|DF_K\|_\infty \leq C(d)\ h_K. \tag{1.11}$$

Consequently, if we introduce the Jacobian of F_K

$$J_K(\widehat{x}) = \det\ DF_K(\widehat{x}), \tag{1.12}$$

we have the estimate

$$\|J_K\|_\infty \leq C\ h_K^d. \tag{1.13}$$

A priori, J_K is a polynomial in $\mathbb{Q}_d$, but this is a not optimal result. A more precise result is given in lemma 1.1.

Lemma 1.1. *In d dimensions, we have* $J_K \in \mathbb{Q}_{d-1}$.

Proof. Let us define :

$$\mathbb{Q}_{r_1,\cdots,r_d} \,=\, span[x_1^{\alpha_1}\cdots\, x_d^{\alpha_d}], \quad \alpha \in \mathcal{S}_{r_1,\cdots,r_d} = \{\alpha \in \mathbb{N}^d, \quad \alpha_j \leq r_j,\ \forall j\}. \tag{1.14}$$

F_K is a $\mathbb{Q}_1$-transformation, so

$$(DF_K)_{i,1} \in \mathbb{Q}_{0,1,\cdots,1}; \quad (DF_K)_{i,2} \in \mathbb{Q}_{1,0,1,\cdots,1}; \quad (DF_K)_{i,d} \in \mathbb{Q}_{1,\cdots,1,0}$$

The determinant of the jacobian matrix DF_K can be written as

$$J_k \,=\, \sum_{\sigma \text{ permutation of } \{1,2,..,d\}} \text{sign}(\sigma)\, (DF_K)_{\sigma(1),1}\ (DF_K)_{\sigma(2),2}\ \cdots\ (DF_K)_{\sigma(d),d}$$

One can recover the degree of J_K, by applying the rule

$$\text{if } p, q \in \mathbb{Q}_{i_1,\cdots,i_d} \times \mathbb{Q}_{j_1,\cdots,j_d},\ \text{ then the product } p\,q \in \mathbb{Q}_{i_1+j_1,\cdots,i_2+j_2}$$

The degree on each variable is equal to :

$$\alpha_i \,=\, 1+\cdots 1+0+1\cdots+1 \,=\, d-1$$

Thus, $J_K \in \mathbb{Q}_{d-1}$. ∎

Remark 1.1. *The reader will notice that the quantities* h_k *and* $diam(K)$ *are somewhat "equivalent" in the sense that*

$$h_K \leq diam(K) \leq C(d)\ \sqrt{d}\, h_K \tag{1.15}$$

To obtain the second inequality, it suffices to remark that

$$diam(K) = \sup\{\ |F_K(\widehat{x}) - F_K(\widehat{y})|,\ (\widehat{x},\widehat{y}) \in \widehat{K}^2\ \}$$

and that, by (1.11)

$$|F_K(\widehat{x}) - F_K(\widehat{y})| \leq C(d)\ h_K\ |\widehat{x} - \widehat{y}|.$$

Next, we observe that by (1.5), $DF_K^{-1}(x)$ is well defined in K and therefore, we can define $\rho_K > 0$ such that:

$$\rho_K{}^{-1} \equiv \|DF_K^{-1}\|_\infty = \sup_{x\in K} |DF_K^{-1}(x)| < +\infty \tag{1.16}$$

Thus, the jacobian J_K^{-1} of DF_K^{-1} satisfies

$$\|J_K^{-1}\|_\infty \leq \frac{C}{\rho_K^d} \tag{1.17}$$

From the fact that $\max\{|F_K^{-1}(x) - F_K^{-1}(y)|\} = \sqrt{d}$, we deduce that

$$\sigma_K \equiv \frac{h_K}{\rho_K} \geq \sqrt{d} \tag{1.18}$$

Let $\mathcal{T}_h = \bigcup K$ be a mesh made of quadrilaterals (or hexahedra) to approximate the geometry of Ω. As usual, we assume that if $\{\mathcal{T}_h\}_h$ is a family of meshes, then

$$\exists\, \sigma > 0 \text{ such that } \forall\, \mathcal{T}_h \in \{\mathcal{T}_h\}_h\ ,\ \forall K \in \mathcal{T}_h,\ \sigma_K < \sigma\ . \tag{1.19}$$

The fact that, for a family of elements K, the "quality" factor σ_K remains bounded will mean that the elements do not degenerate.

More precisely,

- In two dimensions, it is shown in [11], that if T_K^i is the triangle whose vertices are the vertices of K except vertex $n^\circ i$ then:

$$\rho_K = \inf_{i=1}^{4} \rho_K^i, \tag{1.20}$$

  where ρ_K^i is the radius of the largest disk included in T_K^i.
- In three dimensions, we conjecture the following result : let us define a non-degenerate sub-tetrahedron of K as a tetrahedron whose vertices are the image by F_K of four non-coplanar vertices of the reference unit cube $\widehat{K}$. Let $\mathcal{T}_K$ be the set of non-degenerate sub-tetrahedra of K, then:

$$\rho_K = \inf_{T \in \mathcal{T}_K} \rho_T, \tag{1.21}$$

  where ρ_T is the radius of the largest ball included in T.

The property 1.19 implies that we can bound ρ_K independently of K :

$$\rho_K^{-1} \ \leq\ C\, h_K^{-1} \tag{1.22}$$

Using 1.22, we obtain the following estimates:

$$\begin{cases} \|DF_K\|_{\infty,\widehat{K}} \leq C\ h_K \quad ,\ \|J_K\|_{\infty,\widehat{K}} \leq C h_K^d \\ |DF_K^{-1}\|_{\infty,\widehat{K}} \leq C\, h_K^{-1} \ ,\ \|J_K^{-1}\|_{\infty,\widehat{K}} \leq C h_K^{-d} \end{cases} \tag{1.23}$$

B. The finite element space and the semi-discrete problem

We introduce the classical finite element space

$$U_h^r = \left\{ u_h \in C^0(\mathcal{T}_h) \text{ such that } u_{h|_K} \circ F_K \in \mathbb{Q}_r(\widehat{K}) \right\} \subset H^1(\Omega)\ ,$$

As mentioned in the introduction, we want to analyze a scheme which uses Gauss-Lobatto points both for degrees of freedom and quadrature points to enable mass lumping. We shall consider the use of the Gauss-Legendre quadrature formulas to compute the stiffness matrix. For more details about Gauss-Lobatto or Gauss-Legendre points and related quadrature formulas, we refer the reader to [19] or [21]. Let us simply recall they are obtained by tensor products of 1-D points, as illustrated in figure 3. We recall below the fundamental accuracy property of the Gauss and Gauss-Lobatto quadrature rules, that we shall use in the error analysis.

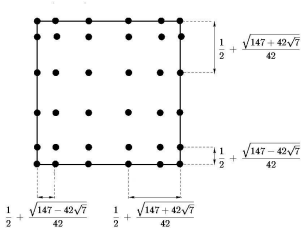


FIG. 3. 2D Gauss-Lobatto points for $r = 5$.

Proposition 1.1. *If we denote by* $(\widehat{\omega}_i^G, \widehat{\xi}_i^G)$ *and* $(\widehat{\omega}_i^{GL}, \widehat{\xi}_i^{GL})$, $1 \leq i \leq (r+1)^d$, *the Gauss and Gauss-Lobatto weights and points in* $\widehat{K}$ *[19], then we have:*

$$\int_{\widehat{K}} \widehat{f}(\widehat{x})\, d\widehat{x} \;=\; \sum_{k=1}^{(r+1)^d} \widehat{\omega}_k^G\; \widehat{f}(\widehat{\xi}_k^G), \quad \forall\, \widehat{f} \in \mathbb{Q}_{2r+1}(\widehat{K})\;, \tag{1.24}$$

$$\int_{\widehat{K}} \widehat{g}(\widehat{x})\, d\widehat{x} \;=\; \sum_{k=1}^{(r+1)^d} \widehat{\omega}_k^{GL}\; \widehat{g}(\widehat{\xi}_k^{GL}), \quad \forall\, \widehat{g} \in \mathbb{Q}_{2r-1}(\widehat{K})\;. \tag{1.25}$$

The weights $\widehat{\omega}_k^G$, $\widehat{\omega}_k^{GL}$ of Gauss and Gauss-Lobatto formulas are strictly positive. In what follows, we use the following notations:

$$\int_{\widehat{K}}^{GL} \widehat{f}(\widehat{x})\, d\widehat{x} \;\stackrel{def}{=}\; \sum_{k=1}^{(r+1)^d} \widehat{\omega}_k^{GL}\; \widehat{f}(\widehat{\xi}_k^{GL}) \qquad \int_{\widehat{K}}^{G} \widehat{f}(\widehat{x})\, d\widehat{x} \;\stackrel{def}{=}\; \sum_{k=1}^{(r+1)^d} \widehat{\omega}_k^{G}\; \widehat{f}(\widehat{\xi}_k^{G})$$

Using the change of variable $x = F_K(\hat{x})$ from K onto $\hat{K}$. Set

$$f(x) \;=\; \hat{f}(F_K^{-1}(x)) \quad \Longrightarrow \quad \int_K f(x)\, dx \;=\; \int_{\widehat{K}} \widehat{f}(\widehat{x})\, |J_K(\widehat{x})|\, d\widehat{x}$$

Thus, we define the approximate integrals on K and Ω as :

$$\int_K^{G,GL} f(x) \;\stackrel{def}{=}\; \sum_{k=1}^{(r+1)^d} \widehat{\omega}_k^{G.GL}\; \widehat{f}(\widehat{\xi}_k^{G,GL})\; |J_K|(\widehat{\xi}_k^G)$$

$$\int_\Omega^{G,GL} f(x) \;\stackrel{def}{=}\; \sum_{K \in \mathcal{T}_h} \int_K^{G,GL} f(x)\;.$$

Then we propose to analyze the following semi-discrete problem :

Find $u_h(t) : [0,T) \longrightarrow U_h^r$ solution of

$$\frac{d^2}{dt^2} \int_\Omega^{GL} u_h(t)\; v_h + \int_\Omega^Q \nabla u_h(t) \cdot \nabla v_h = \int_\Omega f_h(t)\; v_h \quad \forall\, v_h \in U_h^r\;. \tag{1.26}$$

where Q is equal to G or GL, which means that we consider the two cases of using either the Gauss-Legendre or Gauss-Lobatto formulas for the approximate computation of the stiffness matrix.

II. ANALYSIS OF THE METHOD

When an element is a parallelogram or a parallepiped, the transform F_K is affine (in P_1). For a general element K, the transform F_K only belongs to $\mathbb{Q}_1$. For instance, $|J_K|$ is no longer piecewise constant but piecewise polynomial and DF_K^{-1} is now a rational function on each element of the mesh. This is a source of difficulty in the computation and in the analysis, beginning with the interpolation error analysis (see [1]) and influencing the quadrature formulas. Our contribution is clearly related to this last point.

The outline of the rest of the section is the following. Our analysis is essentially based on the error analysis of a family of "elliptic" problems obtained from (1.2) (and for the semi-discrete case from (1.26)) by the application of the Laplace transform (in opposition with [9], who works directly in the time-domain). This is the object of section A.. This analysis is based on an adapted version of the Strang's lemma (see [4] for the usual version of this lemma), namely 2.2, which was already used in [15]. The most delicate part of the work is the analysis of the error due to the numerical integration (cf. theorems 2.6, 2.31 and 2.32). To establish error estimates in the time domain, we use the inverse Laplace-Fourier transform. This is the object of section B..

A. Error estimates for a family of time harmonic elliptic problems

Functional spaces and basic notations

Let Ω be an open set of $\mathbf{R}^d$. We recall that :

$$\begin{aligned} L^2(\Omega) &= \{u \mid \int_\Omega |u|^2 < +\infty\} \\ H^m(\Omega) &= \{u \in L^2(\Omega) \mid \frac{\partial^\alpha u}{\partial x^\alpha} \in L^2(\Omega),\ \forall\, |\alpha| \le m\} \end{aligned} \tag{2.1}$$

Finally, we define the following norms and semi-norms.

$$\begin{aligned} \|u\|_{m,\Omega}^2 &= \sum_{|\alpha|\le m} \int_\Omega \left|\frac{\partial^\alpha u}{\partial x^\alpha}\right|^2 && : \text{the norm in } H^m(\Omega), \\ |u|_{m,\Omega}^2 &= \sum_{|\alpha|= m} \int_\Omega \left|\frac{\partial^\alpha u}{\partial x^\alpha}\right|^2 && : \text{the usual semi-norm in } H^m(\Omega), \\ [u]_{m,\Omega}^2 &= \sum_{i=1}^d \int_\Omega \left|\frac{\partial^m u}{\partial x_i^m}\right|^2 && : \text{the "bracket" semi-norm in } H^m(\Omega). \end{aligned} \tag{2.2}$$

We have obviously the inequality :

$$[u]_{m,\Omega}^2 \le |u|_{m,\Omega}^2 \le \|u\|_{m,\Omega}^2, \quad \forall u \in H^m(\Omega).$$

For functions with complex values and related Sobolev spaces, we use bold symbols. For instance,

$$H^1(\Omega) = \{u \in H^1(\Omega;\mathbb{C})\}.$$

Let $s \in \mathbb{C}^+ = \{s = \eta + i\omega,\ \eta > 0\}$, we introduce the s-dependent inner product norm

$$(u,v)_{1,s} = (\nabla u, \nabla v)_{0,\Omega} + |s|^2 (u,v)_{0,\Omega} \quad , \quad \|u\|_{1,s}^2 = (u,u)_{1,s}$$

Presentation of the problem

In this section, we consider a family of problems parameterized by $s = \eta + i\omega$, when η is fixed, and ω varies in $\mathbb{R}$.

$$\begin{cases} \text{Find } \mathbf{u} : \Omega \longrightarrow \mathbb{C} \text{ solution of} \\ -\Delta \mathbf{u}(x) + s^2 \mathbf{u}(x) = \mathbf{f}(x) \text{ in } \Omega \quad , \quad s \in \mathbb{C}^+ \\ \dfrac{\partial \mathbf{u}}{\partial n}(x) = 0 \quad \text{on } \partial\Omega \end{cases} \tag{2.3}$$

Setting

$$a(s; \mathbf{u}, \mathbf{v}) = (\mathbf{u}, s\mathbf{v})_{1,s} = \bar{s} \left(\int_\Omega s^2 \, u \, \bar{v} + \int_\Omega \nabla \mathbf{u} \cdot \nabla \bar{\mathbf{v}} \right) \quad \text{and} (\mathbf{f}, \mathbf{v}) = \int_\Omega \mathbf{f} \, \bar{\mathbf{v}}, \tag{2.4}$$

the associated variational formulation is :

$$\begin{array}{l} \text{Find } \mathbf{u} \in \mathbf{H}^1(\Omega) \text{ solution of} \\ a(s; \mathbf{u}, \mathbf{v}) = \bar{s}(\mathbf{f}, \mathbf{v}) \quad \forall \, v \in \mathbf{H}^1(\Omega) \ . \end{array} \tag{2.5}$$

We can easily show that this problem satisfies the coercivity relation $\eta \, \|\mathbf{u}\|^2_{1,s} \leq |a(s; \mathbf{u}, \mathbf{u})|$ which insures that it is well-posed in $\mathbf{H}^1(\Omega)$. Then we introduce the following approximate problem (which corresponds to the discretization of (2.5) by spectral finite elements with numerical integration).

$$\begin{array}{l} \text{Find } \mathbf{u}_h \in \mathbf{U}_h^r, \text{ solution of} \\ a_h(s; \mathbf{u}_h, \mathbf{v}_h) = \bar{s}(\mathbf{f}_h, \mathbf{v}_h), \quad \forall \, \mathbf{v}_h \in \mathbf{U}_h^r \ , \end{array} \tag{2.6}$$

where we set $a_h(s; \mathbf{u}_h, \mathbf{v}_h) = \bar{s} \left(\, s^2 \int_\Omega^{GL} \mathbf{u}_h \, \mathbf{v}_h + \int_\Omega^G \nabla \mathbf{u}_h \, \nabla \mathbf{v}_h \, \right)$.

An abstract error estimate The first important property is given by lemma 2.1.

Lemma 2.1. *The bilinear form a_h is uniformly continuous and coercive :*

$$\eta C_1 \, \|\mathbf{v}_h\|^2_{1,s} \leq |a_h(s; \mathbf{v}_h, \mathbf{v}_h)| \qquad \forall \mathbf{v}_h \in \mathbf{U}_h^r$$

$$|a_h(s; \mathbf{u}_h, \mathbf{v}_h)| \leq C_2 |s| \, \|\mathbf{u}_h\|_{1,s} \, \|\mathbf{v}_h\|_{1,s} \quad \forall \mathbf{u}_h, \mathbf{v}_h \in \mathbf{U}_h^r$$

The two constants C_1 and C_2 are independent of h and s.

Proof. These properties are implied by the fact that the constants in the following norm equivalences

$$C_1 \, \|\mathbf{v}_h\|^2_{1,s} \leq \int_K^Q |\nabla \mathbf{u}_h \cdot \nabla \mathbf{u}_h| \leq C_2 \|\mathbf{v}_h\|^2_{1,s}$$

$$C_1 \|\mathbf{v}_h\|_{1,s}^2 \leq \int_K^Q |\mathbf{u}_h \cdot \mathbf{u}_h| \leq C_2 \|\mathbf{v}_h\|_{1,s}^2$$

are independent of h. Let us detail the proof for the stiffness part of a_h (the mass part can be dealt in a similar way). By the definition of the space U_h^r, we have with obvious notations

$$\forall u_h \in U_h^r,\ \forall K \in \tau_h,\ \forall x \in K \quad u_h(x) \,=\, \hat{u}_h(F_K^{-1}(x))$$

and as a consequence for the gradient :

$$\nabla u_h(x) \,=\, (DF_K^*(x))^{-1}\, \hat{\nabla} \hat{u}_h(F_K^{-1}(x)) \tag{2.7}$$

There, if we set $x \,=\, F_K(\hat{x})$, for $v_h \in U_h^r$, we have

$$\nabla u_h(x) \cdot \nabla v_h(x) \,=\, (DF_K(\hat{x}))^{-1}\, (DF_K^*(\hat{x}))^{-1}\, \hat{\nabla} \hat{u}_h(\hat{x}) \cdot \hat{\nabla} \hat{v}_h(\hat{x})$$

By the definition of the approximate integral :

$$\int_\Omega^Q \nabla u_h \cdot \nabla v_h \,=\, \sum_{K \in \tau_h} \int_{\hat{K}}^Q |J_K|\, (DF_K)^{-1}\, (DF_K^*)^{-1}\, \hat{\nabla} \hat{u}_h \cdot \hat{\nabla} \hat{v}_h$$

Thanks to the estimates 1.23, and the positivity of the "Q" quadrature weights :

$$C_1\, h^{d-2} \int_{\hat{K}}^Q |\hat{\nabla} \hat{\mathbf{u}}_h|^2 \,\leq\, \int_{\hat{K}}^Q |J_K|\, DF_K^{-1}\, DF_K^{*-1} \hat{\nabla} \hat{\mathbf{u}}_h \cdot \hat{\nabla} \hat{\mathbf{u}}_h \leq C_2\, h^{d-2} \int_{\hat{K}}^Q |\hat{\nabla} \hat{\mathbf{u}}_h|^2$$

The quadrature points forming an unisolvent set of points in $\mathbb{Q}_r$ and the quadrature weights being strictely positive, the map

$$\hat{u} \mapsto \Big(\int_{\hat{K}}^Q |\hat{u}|^2 \Big)^{1/2}$$

is a norm in $\mathbb{Q}_r$. As all the norms are equivalent in a finite-dimensional space, namely $\mathbb{Q}_r$, we get

$$C_1\, h^{d-2} \int_{\hat{K}} |\hat{\nabla} \hat{\mathbf{u}}_h|^2 \,\leq\, \int_{\hat{K}}^Q |J_K|\, DF_K^{-1}\, DF_K^{*-1} \hat{\nabla} \hat{\mathbf{u}}_h \cdot \hat{\nabla} \hat{\mathbf{u}}_h \leq C_2\, h^{d-2} \int_{\hat{K}} |\hat{\nabla} \hat{\mathbf{u}}_h|^2$$

Transforming back the integrals on $\hat{K}$ into integrals on K, and using again the estimates 1.23, we obtain :

$$C_1 \int_K |\nabla \mathbf{u}_h|^2 \,\leq\, \int_K^Q |\nabla \mathbf{u}_h \cdot \nabla \mathbf{u}_h| \leq C_2 \int_K |\nabla \mathbf{u}_h|^2$$

■

A similar reasoning can be done to get inequalities for the mass term $\int_K^Q |\mathbf{u}_h \cdot \mathbf{u}_h|$. By linear combination, the claimed estimates are obtained. Thanks to lemma 2.1 and following [16], one can prove the following version of Strang's lemma. The proof is classical (see [4]), but we pay attention to the dependency of the constants with respect to $|s|$ and $\eta \,=\, Re(s)$.

Lemma 2.2. *If* $\mathbf{u}$ *is the solution of (2.5) and* $\mathbf{u}_h$ *the solution of (2.6), there exists a constant* $C > 0$ *which does not depend on the step* h *and the parameter* s *such that*

$$\|\mathbf{u}-\mathbf{u}_h\|_{1,s} \leq \underbrace{C \inf_{\mathbf{v}_h \in \mathbf{U}_h^r} \left[\left(1+\frac{|s|}{\eta}\right) \|\mathbf{u}-\mathbf{v}_h\|_{1,s} \right.}_{\text{interpolation error}} + \underbrace{\frac{1}{\eta} \sup_{\mathbf{w}_h \in \mathbf{U}_h^r} \frac{|(a-a_h)(s;\mathbf{v}_h,\mathbf{w}_h)|}{\|\mathbf{w}_h\|_{1,s}}}_{\text{numerical integration error}} \left. \right] . \tag{2.8}$$

In the following, we study separately the two terms of the right hand-side of the Strang's Lemma, namely the interpolation error and the quadrature error.

Results related to the interpolation error:

The following results, related to interpolation error, are particular cases of more general estimates, which are given without proof in [11], in the case two dimensions. We have checked carefully that these results were also valid in 3-D. However we have chosen not to present the proofs here, since this is not the central point of the paper. The interpolation estimates are a consequence of the lemmas 2.3 and 2.4.

Lemma 2.3. *(Bramble-Hilbert)*
Let $r \geq 1,\ m \leq r+1 \Rightarrow \mathbf{H}^{r+1}(\widehat{K}) \subset \mathbf{H}^m(\widehat{K})$. *We denote by* $\widehat{\pi}^r$ *the continuous linear application from* $\mathbf{H}^{r+1}(\widehat{K})$ *into* $\mathbf{H}^m(\widehat{K})$ *satisfying*

$$\forall\, \widehat{\mathbf{p}} \in \mathbb{Q}_r(\widehat{K})\ ,\ \widehat{\pi}^r \widehat{\mathbf{p}} = \widehat{\mathbf{p}}\ .$$

Then, there exists a constant $C > 0$ *which only depends on* $\widehat{K}$ *and* r *such that*

$$\forall\, \widehat{\mathbf{u}} \in \mathbf{H}^{r+1}(\widehat{K}) \quad , \quad \|\widehat{\mathbf{u}} - \widehat{\pi}^r \widehat{\mathbf{u}}\|_{m,\widehat{K}} \leq C\, [\widehat{\mathbf{u}}]_{r+1,\widehat{K}}\ . \tag{2.9}$$

Lemma 2.4. *Let* $K \in \mathcal{T}_h,\ m \geq 1$. *Then* $\forall\, \mathbf{v} \in \mathbf{H}^m(K)$ *there exists* $C_\sigma > 0$ *such that*

$$\|\mathbf{v}\|_{0,K} \leq C_\sigma\, h_K^{\frac{d}{2}}\, \|\widehat{\mathbf{v}}\|_{0,\widehat{K}}\ , \tag{2.10}$$

$$|\mathbf{v}|_{m,K} \leq C_\sigma\, h_K^{\frac{d}{2}-m}\, \|\widehat{\mathbf{v}}\|_{m,\widehat{K}}\ , \tag{2.11}$$

$$[\widehat{\mathbf{v}}]_{m,\widehat{K}} \leq C_\sigma\, h_K^{m-\frac{d}{2}}\, |\mathbf{v}|_{m,K}\ . \tag{2.12}$$

Remark 2.1. *The inequalities 2.10 and 2.11 are obtained, by a change of variables, from* K *to* $\hat{K}$. *For the inequality 2.12, we use the following property, which is true both in 2-D and 3-D :*

$$\frac{\partial^2 F_K}{\partial x_m^2} = 0$$

That is why it is so crucial to use the "bracket" semi-norm $[v]$, *whereas the usual semi-norm is used to obtain similar estimations for the triangular/tetrahedral elements.*

Lemma 2.3 and 2.4 lead to the following result:

Proposition 2.1. *Let $K \in \mathcal{T}_h$ and $0 \leq m \leq l \leq r+1$. We denote by π_K^r the continous linear operator from $\mathbf{H}^l(K)$ into $\mathbf{H}^m(K)$ given by*

$$(\pi_K^r \mathbf{v}) \circ F_K = \widehat{\pi}^r \widehat{\mathbf{v}} \ .$$

Then there exists $C_\sigma > 0$ such that $\forall\ \mathbf{v} \in \mathbf{H}^l(K)$,

$$|\mathbf{v} - \pi_K^r \mathbf{v}|_{m,K} \leq C_\sigma \ h_K^{l-m} \ |\mathbf{v}|_{l,K} \ . \tag{2.13}$$

This yields the following interpolation error theorem

Theorem 2.5. *If $\mathbf{u}$ is the solutions of (2.5) and if $\mathbf{f}$ is smooth enough such that $\mathbf{u} \in \mathbf{H}^{r+1}(\Omega)$, then there exists $C > 0$ such that*

$$\inf_{\mathbf{v}_h \in \mathbf{U}_h^r} \left[\ \left(1 + \frac{|s|}{\eta}\right) \|\mathbf{u} - \mathbf{v}_h\|_{1,s} \right] \leq C \ \left(1 + \frac{|s|}{\eta}\right) \left(\ |s||\mathbf{u}|_{r,\Omega} + |\mathbf{u}|_{r+1,\Omega} \ \right) h^r \ . \tag{2.14}$$

Towards the estimations of the quadrature error.

We decompose the numerical integration error of Lemma 2.2, into a sum of two terms, that we will analyze separately in the next two subsections.

$$\sup_{\mathbf{w}_h \in \mathbf{U}_h^r} |(a - a_h)(s; \mathbf{v}_h, \mathbf{w}_h)| \leq |s|^3 \sup_{\mathbf{w}_h \in \mathbf{U}_h^r} e_h^m(\mathbf{v}_h, \mathbf{w}_h) + |s| \sup_{\mathbf{w}_h \in \mathbf{U}_h^r} e_h^s(\mathbf{v}_h, \mathbf{w}_h) \tag{2.15}$$

where $e_h^m(\mathbf{v}_h, \mathbf{w}_h)$ and $e_h^s(\mathbf{v}_h, \mathbf{w}_h)$ are quadrature error terms, respectively associated with the mass matrix and stiffness matrix :

$$e_h^m(\mathbf{v}_h, \mathbf{w}_h) = \left| \int_\Omega \mathbf{v}_h \ \bar{\mathbf{w}}_h - \int_\Omega^{GL} \mathbf{v}_h \ \bar{\mathbf{w}}_h \right| \ ,$$

$$e_h^s(\mathbf{v}_h, \mathbf{w}_h) = \left| \int_\Omega \nabla \mathbf{v}_h \cdot \nabla \bar{\mathbf{w}}_h - \int_\Omega^{Q} \nabla \mathbf{v}_h \cdot \nabla \bar{\mathbf{w}}_h \right| \ , \quad Q = G \text{ or } GL.$$

Influence of numerical integration for the mass term: estimate of e_h^m.

For any functions $(\mathbf{v}_h, \mathbf{w}_h) \in \mathbf{U}_h^r \times \mathbf{U}_h^r$ we have $\displaystyle\int_K \mathbf{v}_h \ \bar{\mathbf{w}}_h - \int_K^{GL} \mathbf{v}_h \ \bar{\mathbf{w}}_h = \widehat{E}_K(\widehat{\mathbf{v}}_h, \widehat{\mathbf{w}}_h)$ where $\widehat{E}_K(\widehat{\mathbf{v}}_h, \widehat{\mathbf{w}}_h) = \displaystyle\int_{\widehat{K}} |J_K| \ \widehat{\mathbf{v}}_h \ \widehat{\bar{\mathbf{w}}}_h - \int_{\widehat{K}}^{GL} |J_K| \ \widehat{\mathbf{v}}_h \ \widehat{\bar{\mathbf{w}}}_h$ with $\widehat{\mathbf{v}}_h \overset{def}{=} \mathbf{v}_h \circ F_K$.

Proposition 2.2. *Let $j \geq 1$ and $\widehat{\Pi}_j$ the projector from $L^2(\widehat{K})$ into $\mathbb{Q}_j(\widehat{K})$. Then*

$$\forall\, r \geq d, \quad \forall\, p \in [1, r-d+1], \quad \forall\, (\mathbf{v}_h, \mathbf{w}_h) \in \mathbf{U}_h^r \times \mathbf{U}_h^r \ ,$$
$$\widehat{E}_K(\ \widehat{\mathbf{v}}_h, \widehat{\mathbf{w}}_h\ ) = \widehat{E}_K(\ \widehat{\mathbf{v}}_h - \widehat{\Pi}_{p-1}\widehat{\mathbf{v}}_h\ ,\ \widehat{\mathbf{w}}_h - \widehat{\Pi}_0\widehat{\mathbf{w}}_h\ )\ .$$

Proof. For any functions $(\widehat{\mathbf{v}}_h, \widehat{\mathbf{w}}_h) \in \mathbf{U}_h^r \times \mathbf{U}_h^r$ we have

$$\begin{aligned}\widehat{E}_K(\ \widehat{\mathbf{v}}_h - \widehat{\Pi}_{p-1}\widehat{\mathbf{v}}_h\ ,\ \widehat{\mathbf{w}}_h - \widehat{\Pi}_0\widehat{\mathbf{w}}_h\ ) &= \widehat{E}_K(\ \widehat{\mathbf{v}}_h\ ,\ \widehat{\mathbf{w}}_h\ ) - \widehat{E}_K(\ \widehat{\Pi}_{p-1}\widehat{\mathbf{v}}_h\ ,\ \widehat{\mathbf{w}}_h\ ) \\ &\quad - \widehat{E}_K(\ \widehat{\mathbf{v}}_h\ ,\ \widehat{\Pi}_0\widehat{\mathbf{w}}_h\ ) + \widehat{E}_K(\ \widehat{\Pi}_{p-1}\widehat{\mathbf{v}}_h\ ,\ \widehat{\Pi}_0\widehat{\mathbf{w}}_h\ )\ .\end{aligned}$$

So, the lemma is true if

$$\widehat{E}_K(\ \widehat{\Pi}_{p-1}\widehat{\mathbf{v}}_h\ ,\ \widehat{\mathbf{w}}_h\ ) = \widehat{E}_K(\ \widehat{\mathbf{v}}_h\ ,\ \widehat{\Pi}_0\widehat{\mathbf{w}}_h\ ) = \widehat{E}_K(\ \widehat{\Pi}_{p-1}\widehat{\mathbf{v}}_h\ ,\ \widehat{\Pi}_0\widehat{\mathbf{w}}_h\ ) = 0\ .$$

According to (1.25), $\widehat{E}_K(\ \widehat{\Pi}_{p-1}\widehat{\mathbf{v}}_h\ ,\ \widehat{\mathbf{w}}_h\ ) = 0$ if $\deg(\ |J_K|\ (\widehat{\Pi}_{p-1}\widehat{\mathbf{v}}_h)\ \widehat{\mathbf{w}}_h\ ) \leq 2r-1$.

As $J_K \in \mathbb{Q}_{d-1}$, $\widehat{\Pi}_{p-1}\widehat{\mathbf{v}}_h \in \mathbb{Q}_{p-1}$ and $\widehat{\mathbf{w}}_h \in V_h^r$, we have

$$|J_K|\ (\widehat{\Pi}_{p-1}\widehat{\mathbf{v}}_h)\ \widehat{\mathbf{w}}_h \in \mathbb{Q}_{p+d+r-2}\ .$$

Consequently,

$$\widehat{E}_K(\ \widehat{\Pi}_{p-1}\widehat{\mathbf{v}}_h\ ,\ \widehat{\mathbf{w}}_h\ ) = 0 \ \text{ if } \ p+r+d-2 \leq 2r-1 \Leftrightarrow p \leq r-d+1\ . \tag{2.16}$$

By the same kind of arguments, as $|J_K|\ (\widehat{\Pi}_0\widehat{\mathbf{w}}_h)\ \widehat{\mathbf{v}}_h \in \mathbb{Q}_{d+r-1}$ then

$$\widehat{E}_K(\ \widehat{\mathbf{v}}_h\ ,\ \widehat{\Pi}_0\widehat{\mathbf{w}}_h\ ) = 0 \ \text{ if } \ d+r-1 \leq 2r-1 \Leftrightarrow r \geq d\ . \tag{2.17}$$

Finally, as $|J_K|\ (\widehat{\Pi}_{p-1}\widehat{\mathbf{v}}_h)\ (\widehat{\Pi}_0\widehat{\mathbf{w}}_h) \in \mathbb{Q}_{d+p-2}$ then

$$\widehat{E}_K(\ \widehat{\Pi}_{p-1}\widehat{\mathbf{v}}_h\ ,\ \widehat{\Pi}_0\widehat{\mathbf{w}}_h\ ) = 0 \ \text{ if } \ p+d-2 \leq 2r-1 \Leftrightarrow p \leq 2r-d+1\ . \tag{2.18}$$

The proposition is thus demonstrated thanks to (2.16) and (2.17). ∎

Proposition 2.3. *Let $r \geq d$. $\exists\, C > 0$ such that $\forall\, (\mathbf{v}_h, \mathbf{w}_h) \in \mathbf{U}_h^r \times \mathbf{U}_h^r$.*

$$\left| \int_K \mathbf{v}_h\ \bar{\mathbf{w}}_h - \int_K^{GL} \mathbf{v}_h\ \bar{\mathbf{w}}_h \right| \leq C\ h_K^{p+1}\ |\mathbf{v}_h|_{p,K}\ |\mathbf{w}_h|_{1,K} \quad , \quad \forall p\ \in [1; r-d+1]\ . \tag{2.19}$$

Proof. Using Cauchy-Schwartz inequality and the previous proposition we have:

$$\begin{aligned}|\widehat{E}_K(\ \widehat{\mathbf{v}}_h, \widehat{\mathbf{w}}_h\ )| &\leq \|J_K\|_{\infty,\widehat{K}}\ \|\widehat{\mathbf{v}}_h - \widehat{\Pi}_{p-1}\widehat{\mathbf{v}}_h\|_{0,\widehat{K}}\ \|\widehat{\mathbf{w}}_h - \widehat{\Pi}_0\widehat{\mathbf{w}}_h\|_{0,\widehat{K}} \\ &\leq C\|J_K\|_{\infty,\widehat{K}}\ [\widehat{\mathbf{v}}_h]_{p,\widehat{K}}\ [\widehat{\mathbf{w}}_h]_{1,\widehat{K}},\ \text{according to (2.9)} \\ &\leq C\|J_K\|_{\infty,\widehat{K}}\ h_K^{p-\frac{d}{2}}\ |\mathbf{v}_h|_{p,K}\ h_K^{1-\frac{d}{2}}|\mathbf{w}_h|_{1,K},\ \text{according to (2.12)} \\ &\leq C\ h_K^{p+1}|\mathbf{v}_h|_{p,K}\ |\mathbf{w}_h|_{1,K},\ \text{by (1.23)}\end{aligned}$$

∎

Now, as $|\mathbf{w}_h|_{1,K} < \|\mathbf{w}_h\|_{1,s}$, using the result of the previous proposition with $p = r-d+1$ and taking the sum over the elements of the mesh, we can state the following theorem:

Theorem 2.6. *Let $r \geq d$. $\exists\, C > 0$ such that $\forall\ (\mathbf{v}_h, \mathbf{w}_h) \in \mathbf{U}_h^r \times \mathbf{U}_h^r$,*

$$e_h^m(\mathbf{v}_h, \mathbf{w}_h) \leq C\ h^{r-d+2}\ |\mathbf{v}_h|_{r-1,\Omega}\ \|\mathbf{w}_h\|_{1,s}\ . \tag{2.20}$$

Remark 2.2. *The condition $r \geq d$ is a bit surprising. It is in fact a pure algebraic limitation of our proof. Numerical results in section III show that for $r < d$ we have exactly the same kind of convergence.*

Remark 2.3. *In the case of a general mesh, the space dimension will have an influence on the final error estimate. Indeed, in two dimensions, we see that we get $h^{r-d+2} = h^r$, which gives the optimal H^1 convergence rate for the method without numerical integration. However, in three dimensions, $h^{r-d+2} = h^{r-1}$, which means that we will lose a priori one order of convergence, that is confirmed by the further numerical results.*

Remark 2.4. *In the case of a mesh made of parallelograms or parallepipeds, there is no loss of order even for $d = 3$. Indeed, $|J_K| = h_K^d$ is piecewise constant (on each element). Then relations (2.16), (2.17) and (2.18) become respectively $(p \leq r)$, $(1 \leq r)$, $(p \leq 2r)$ and thus, no longer depend on d. In this case we simply have*

$$e_h^m(\mathbf{v}_h, \mathbf{w}_h) \leq C\ h^{r+1}\ |\mathbf{v}_h|_{r-1,\Omega}\ \|\mathbf{w}_h\|_{1,s} \quad \forall (\mathbf{v}_h, \mathbf{w}_h) \in \mathbf{U}_h^r \times \mathbf{U}_h^r\ . \tag{2.21}$$

Influence of numerical integration for the stiffness term: estimate of e_h^s.

We cannot use the same arguments for the e_h^m, due to the fact that numerical integration is performed on terms which contain the quantity DF_K^{-1} which is not polynomial but a rational function (in the case of a general mesh).

$$e_h^s(v_h, w_h) \ = \ \sum_{K \in \tau_h} E_K^s(v_h, w_h)$$

where

$$E_K^s(v_h, w_h) \ = \ \int_K \nabla \mathbf{v}_h \cdot \nabla \bar{\mathbf{w}}_h \ - \ \int_K^Q \nabla \mathbf{v}_h \cdot \nabla \bar{\mathbf{w}}_h$$

Setting $v_h(x) \ = \ \hat{v}_h(F_K^{-1}(x))\ w_h(x) \ = \ \hat{w}_h(F_K^{-1}(x))$ and proceeding as in the proof of lemma 2.1, we have

$$E_K^s(v_h, w_h) \ = \ \int_{\widehat{K}} |J_K|\ DF_K^{-1} DF_K^{*-1}\, \widehat{\nabla} \widehat{\mathbf{v}}_h \cdot \widehat{\nabla} \widehat{\bar{\mathbf{w}}}_h \ - \ \int_{\widehat{K}}^Q |J_K|\ DF_K^{-1} DF_K^{*-1} \widehat{\nabla} \widehat{\mathbf{v}}_h \cdot \widehat{\nabla} \widehat{\bar{\mathbf{w}}}_h \tag{2.22}$$

Remark 2.5. *When the mesh is uniform, the term $|J_K|\ DF_K^{-1} DF_K^{*-1}$ is proportional to the identity matrix for any element K. According to (1.24), the use of the Gauss quadrature rule to compute the stiffness matrix is equivalent to exact integration because $\widehat{\nabla} \widehat{\mathbf{v}}_h \cdot \widehat{\nabla} \widehat{\bar{\mathbf{w}}}_h$ lies in $\mathbb{Q}_{2r}(\widehat{K})$. So, the use of Gauss quadrature rule has no influence on the accuracy of the method.*

For the error analysis, it is more convenient to use again the relation 2.7 in the reverse

way for w_h so that 2.22 becomes

$$E_K^s(v_h, w_h) \;=\; \int_{\widehat{K}} |J_K|\; DF_K^{-1}\nabla \mathbf{v}_h \cdot \widehat{\nabla}\widehat{\mathbf{w}}_h \;-\; \int_{\widehat{K}}^{Q} |J_K|\; DF_K^{-1}\nabla \mathbf{v}_h \cdot \widehat{\nabla}\widehat{\mathbf{w}}_h \tag{2.23}$$

Proposition 2.4. *Assuming that $Q = G$, there exists a constant $C > 0$ such that*

$$E_K^s(v_h, w_h) \leq C\; h_K^{r+1} \left(\sum_{m=0}^{d-1} h_K^{-m}\; |\mathbf{v}_h|_{r+2-m,K} \right) |\mathbf{w}|_{1,K} \tag{2.24}$$

Proof.

Let $\widehat{\mathcal{J}}^r$ the $\mathbb{Q}_r$ interpolation operator on Gauss points. For any $\widehat{\varphi}$, continuous in $\hat{K}$:

$$\widehat{\mathcal{J}}^r(\widehat{\varphi}) \in \mathbb{Q}_r(\widehat{K}) \quad \text{et} \quad \widehat{\varphi}(\widehat{\xi}_i^G) = \widehat{\mathcal{J}}^r(\widehat{\varphi})(\widehat{\xi}_i^G)\;,\; \forall\; 1 \leq i \leq (r+1)^d\;.$$

Setting $\widehat{\underline{\mathbf{z}}}_h^K = |J_K|\; DF_K^{-1}\; \nabla \mathbf{v}_h$ we can write:

$$\begin{aligned} E_K^s(v_h, w_h) &= \int_{\widehat{K}} \widehat{\underline{\mathbf{z}}}_h^K \cdot \widehat{\nabla}\widehat{\mathbf{w}}_h - \sum_{i=1}^{(r+1)^d} \widehat{\omega}_i^G\; \widehat{\underline{\mathbf{z}}}_h^K(\widehat{\xi}_i^G) \cdot \widehat{\nabla}\widehat{\mathbf{w}}_h(\widehat{\xi}_i^G) \\ &= \int_{\widehat{K}} \widehat{\underline{\mathbf{z}}}_h^K \cdot \widehat{\nabla}\widehat{\mathbf{w}}_h - \sum_{i=1}^{(r+1)^d} \widehat{\omega}_i^G\; \widehat{\mathcal{J}}^r(\widehat{\underline{\mathbf{z}}}_h^K)(\widehat{\xi}_i^G) \cdot \widehat{\nabla}\widehat{\mathbf{w}}_h(\widehat{\xi}_i^G) \\ &= \int_{\widehat{K}} \widehat{\underline{\mathbf{z}}}_h^K \cdot \widehat{\nabla}\widehat{\mathbf{w}}_h - \int_{\widehat{K}}^{G} \widehat{\mathcal{J}}^r(\widehat{\underline{\mathbf{z}}}_h^K) \cdot \widehat{\nabla}\widehat{\mathbf{w}}_h \end{aligned} \tag{2.25}$$

Since $\widehat{\mathcal{J}}^r(\widehat{\underline{\mathbf{z}}}_h^K) \cdot \widehat{\nabla}\widehat{\mathbf{w}}_h \in \mathbb{Q}_{2r}(\widehat{K})$, we use the accuracy property of the Gauss-Legendre formulas, namely (1.24) to rewrite $E_K^s(v_h, w_h)$ as an integral and get

$$E_K^s(v_h, w_h) \;=\; \int_{\widehat{K}} (\widehat{\underline{\mathbf{z}}}_h^K - \widehat{\mathcal{J}}^r(\widehat{\underline{\mathbf{z}}}_h^K)) \cdot \widehat{\nabla}\widehat{\mathbf{w}}_h \;=\; \|\widehat{\underline{\mathbf{z}}}_h^K - \widehat{\mathcal{J}}^r(\widehat{\underline{\mathbf{z}}}_h^K)\|_{0,\widehat{K}}\; |\widehat{\mathbf{w}}_h|_{1,\widehat{K}}\;. \tag{2.26}$$

To bound the right hand side, we use the Bramble-Hilbert lemma 2.3.

$$\|\widehat{\underline{\mathbf{z}}}_h^K - \widehat{\mathcal{J}}^r(\widehat{\underline{\mathbf{z}}}_h^K)\|_{0,\widehat{K}} \leq [\widehat{\underline{\mathbf{z}}}_h^K]_{r+1,\widehat{K}}\;.$$

To estimate the term $[\widehat{\underline{\mathbf{z}}}_h^K]_{r+1,\widehat{K}}$, the main difficulty is to estimate the derivatives of the rational function

$$\widehat{\underline{\mathbf{z}}}_h^K \;=\; |J_K|\; DF_K^{-1}\nabla v_h.$$

For this, we note that the matrix

$$M_K = |J_K|\; DF_K^{-1} \quad (:= (m_{ij}^K))$$

is nothing but than the transpose of the cofactor matrix of DF_K, one realizes that :

$$\forall\; (i,j) = \{1,..,d\}^2,\; m_{ij}^K \in \mathbb{Q}_{d-1}(\widehat{K})\;. \tag{2.27}$$

Using the Leibniz formula we have:

$$[\underline{\widehat{\mathbf{z}}}_h^K]^2_{r+1,\widehat{K}} \leq \sum_{l=1}^{d}\sum_{i=1}^{d}\sum_{j=1}^{d}\int_{\widehat{K}} | \frac{\partial^{r+1}}{\partial \widehat{x}_l^{r+1}}(m_{ij}^K(\nabla \mathbf{v}_h \circ F_K)_j) |^2$$

$$\leq \sum_{l=1}^{d}\sum_{i=1}^{d}\sum_{j=1}^{d}\sum_{m=0}^{r+1}\int_{\widehat{K}} \begin{pmatrix} r+1 \\ m \end{pmatrix} \left| \frac{\partial^m}{\partial \widehat{x}_l^m}(m_{ij}^K) \frac{\partial^{r+1-m}}{\partial \widehat{x}_l^{r+1-m}}(\nabla \mathbf{v}_h \circ F_K)_j \right|^2$$

According to (2.27),

$$\frac{\partial^m}{\partial \widehat{x}_l^m}(m_{ij}^K) = 0 \quad \text{if } m \geq d.$$

Moreover it is shown in [18], that

$$| \frac{\partial^m}{\partial \widehat{x}_l^m}(m_{ij}^K) | < C\ h_K^{d-1}. \tag{2.28}$$

We deduce from (2.28) that there exists $C > 0$ such that

$$[\underline{\widehat{\mathbf{z}}}_h^K]^2_{r+1,\widehat{K}} \leq C\ h_K^{2(d-1)} \sum_{l=1}^{d}\sum_{i=1}^{d}\sum_{j=1}^{d}\sum_{m=0}^{d-1}\int_{\widehat{K}} | \frac{\partial^{r+1-m}}{\partial \widehat{x}_l^{r+1-m}}(\nabla \mathbf{v}_h \circ F_K)_j |^2$$

$$= C\ h_K^{2(d-1)} \sum_{m=0}^{d-1} [\nabla \mathbf{v}_h \circ F_K]^2_{r+1-m,\widehat{K}}$$

Thus, using lemma 2.4 and more precisely (2.12), we get

$$[\underline{\widehat{\mathbf{z}}}_h^K]^2_{r+1,\widehat{K}} \leq C\ h_K^{2(d-1)} \sum_{m=0}^{d-1} h^{2(r+1-m-\frac{d}{2})}\ |\mathbf{v}_h|^2_{r+2-m,K}$$

$$\leq C\ h_K^{2r+d} \sum_{m=0}^{d-1} h_K^{-2m}\ |\mathbf{v}_h|^2_{r+2-m,K}\ ,$$

that is to say

$$[\underline{\widehat{\mathbf{z}}}_h^K]_{r+1,\widehat{K}} \leq C\ h_K^{r+\frac{d}{2}} \sum_{m=0}^{d-1} h_K^{-m}\ |\mathbf{v}_h|_{r+2-m,K}. \tag{2.29}$$

It remains to estimate $|\widehat{\mathbf{w}}_h|_{1,\widehat{K}}$ in (2.26). Using successively (1.23) and (1.19), we have

$$|\widehat{\mathbf{w}}_h|^2_{1,\widehat{K}} = \int_{\widehat{K}} |\widehat{\nabla}\widehat{\mathbf{w}}_h|^2\ d\widehat{x} = \int_K |DF_K^* \nabla \mathbf{w}_h|^2\ |J_K|^{-1}\ dx \quad (x = F_K(\widehat{x}))$$

$$\leq \|DF_K\|^2_{\infty,\widehat{K}}\ \|J_K^{-1}\|_{\infty,\widehat{K}}\ |\mathbf{w}_h|^2_{1,K}$$

$$\leq C\ \frac{h_K^2}{\rho_K^d}\ |\mathbf{w}_h|^2_{1,K}\ \leq C\ \sigma\ h_K^{1-\frac{d}{2}}\ |\mathbf{w}_h|_{1,K}\ . \tag{2.30}$$

Finally, substituting (2.29) and (2.30) into (2.26) leads to the result. ∎

Using the last proposition and summing over the elements of the mesh, we can estimate the error coming from the use of a Gauss quadrature formula to compute the stiffness matrix.

Theorem 2.7. *Assuming $Q = G$, there exists $C > 0$ such that $\forall\ (\mathbf{v}_h, \mathbf{w}_h) \in \mathbf{U}_h^r \times \mathbf{U}_h^r$,*

$$e_h^s(\mathbf{v}_h, \mathbf{w}_h) \leq C\ h^{r+1}\ \left(\sum_{m=0}^{d-1} h^{-m}\ |\mathbf{v}_h|_{r+2-m,\Omega} \right)\ \|\mathbf{w}_h\|_{1,s}\ . \tag{2.31}$$

Our next result concerns the case where one uses the Gauss-Lobatto formula to evaluate the stiffness matrix. The equivalent of theorem 2.7 is:

Theorem 2.8. *Assuming $Q = GL$, there exists $C > 0$ such that $\forall\ (\mathbf{v}_h, \mathbf{w}_h) \in \mathbf{U}_h^r \times \mathbf{U}_h^r$,*

$$e_h^s(\mathbf{v}_h, \mathbf{w}_h) \leq C\ h^{r}\ \left(\sum_{m=0}^{d-1} h^{-m}\ (|\mathbf{v}_h|_{r+2-m,\Omega} + |\mathbf{v}_h|_{r+1-m,\Omega}) \right)\ \|\mathbf{w}_h\|_{1,s}\ . \tag{2.32}$$

Proof. We simply indicate how to modify the proof of proposition. The rest of the proof is straightforward. In this proof, $\widehat{\mathcal{J}}^r$ still denotes the $\mathbb{Q}_r$-interpolation operator on Gauss-Lobatto points while $\widehat{\mathcal{J}}^{r-1}$ the $\mathbb{Q}_{r-1}$ interpolation operator on a $\mathbb{Q}_{r-1}$ unisolvent points.

$$\begin{aligned} e_h^s(\mathbf{v}_h, \mathbf{w}_h) &= \int_{\widehat{K}} \widehat{\underline{\mathbf{z}}}_h^K \cdot \widehat{\nabla}\widehat{\mathbf{w}}_h - \int_{\widehat{K}}^{GL} \underbrace{\widehat{\mathcal{J}}^{r-1}(\widehat{\underline{\mathbf{z}}}_h^K) \cdot \widehat{\nabla}\widehat{\mathbf{w}}_h}_{\in\, \mathbb{Q}_{2r-1}(\widehat{K})} - \int_{\widehat{K}}^{GL} (\widehat{\mathcal{J}}^r(\widehat{\underline{\mathbf{z}}}_h^K)\ -\ \widehat{\mathcal{J}}^{r-1}(\widehat{\underline{\mathbf{z}}}_h^K)) \cdot \widehat{\nabla}\widehat{\mathbf{w}}_h \\ &= \int_{\widehat{K}} (\widehat{\underline{\mathbf{z}}}_h^K - \widehat{\mathcal{J}}^{r-1}(\widehat{\underline{\mathbf{z}}}_h^K)) \cdot \widehat{\nabla}\widehat{\mathbf{w}}_h\ -\ \int_{\widehat{K}}^{GL} (\widehat{\mathcal{J}}^r(\widehat{\underline{\mathbf{z}}}_h^K)\ -\ \widehat{\mathcal{J}}^{r-1}(\widehat{\underline{\mathbf{z}}}_h^K)) \cdot \widehat{\nabla}\widehat{\mathbf{w}}_h \end{aligned} \tag{2.33}$$

The first term can be dealt in a similar way than for the previous proof.

$$|\int_{\widehat{K}} (\widehat{\underline{\mathbf{z}}}_h^K - \widehat{\mathcal{J}}^{r-1}(\widehat{\underline{\mathbf{z}}}_h^K)) \cdot \widehat{\nabla}\widehat{\mathbf{w}}_h| \leq C\ h^r\ \left(\sum_{m=0}^{d-1} h^{-m}\ |\mathbf{v}_h|_{r+1-m,\Omega} \right)\ \|\mathbf{w}_h\|_{1,s}$$

For the second term, we use the equivalence between the L^2 norm and GL-norm

$$\begin{aligned} &|\textstyle\int_{\widehat{K}}^{GL} (\widehat{\mathcal{J}}^r(\widehat{\underline{\mathbf{z}}}_h^K)\ -\ \widehat{\mathcal{J}}^{r-1}(\widehat{\underline{\mathbf{z}}}_h^K)) \cdot \widehat{\nabla}\widehat{\mathbf{w}}_h| \\ &\qquad \leq C \Big(\int_{\widehat{K}}^{GL} |\widehat{\mathcal{J}}^r(\widehat{\underline{\mathbf{z}}}_h^K)\ -\ \widehat{\mathcal{J}}^{r-1}(\widehat{\underline{\mathbf{z}}}_h^K)|^2 \Big)^{1/2} \Big(\int_{\widehat{K}}^{GL} |\widehat{\nabla}\widehat{\mathbf{w}}_h|^2 \Big)^{1/2} \\ &\qquad \leq C\, \|(\widehat{\mathcal{J}}^r(\widehat{\underline{\mathbf{z}}}_h^K)\ -\ \widehat{\mathcal{J}}^{r-1}(\widehat{\underline{\mathbf{z}}}_h^K))\|_{0,\hat{K}}\, |\widehat{\mathbf{w}}_h|_{1,\widehat{K}} \\ &\qquad \leq C\, (\|\widehat{\mathcal{J}}^r(\widehat{\underline{\mathbf{z}}}_h^K)\ -\ \widehat{\underline{\mathbf{z}}}_h^K\|_{0,\hat{K}}\ +\ \|\widehat{\mathcal{J}}^{r-1}(\widehat{\underline{\mathbf{z}}}_h^K)\ -\ \widehat{\underline{\mathbf{z}}}_h^K\|_{0,\hat{K}})\, |\widehat{\mathbf{w}}_h|_{1,\widehat{K}} \end{aligned} \tag{2.34}$$

The first term is equal to the previous error computed with Gauss points. By summing the differents estimates, we obtain the expected result. ∎

Remark 2.6. *In general, one can expect from theorems 2.31 and 2.32, to lose one order of accuracy by using Gauss-Lobatto formulas instead of Gauss formulas for the stiffness matrix. This is confirmed by numerical experiments in 3-D.*

Remark 2.7. *In the case of mesh made of parallelograms/parallepipeds, the cofactors* m_{ij}^K *are constant by element. Then the relation 2.32 becomes*

$$e_h^s(\mathbf{v}_h, \mathbf{w}_h) \leq C\ h^r(|\mathbf{v}_h|_{r+2,\Omega} + |\mathbf{v}_h|_{r+1,\Omega}) \tag{2.35}$$

The use of Gauss-Lobatto rules for stiffness matrix does not deteriorate the order of convergence, in the case of an uniform mesh.

Global error estimates

We introduce a particular semi-norm, for $m \geq 1$

$$|||v_h|||_{m,\Omega} = |v_h|_{m,\Omega} + |v_h|_{m-1,\Omega}$$

Using the decomposition 2.15 with theorems 2.6 and 2.31 (or 2.32), one easily obtains the following lemma

Lemma 2.9. *We have the following estimate:*

$$\sup_{\mathbf{w}_h \in \mathbf{U}_h^r} \frac{|(a-a_h)(s;\mathbf{v}_h,\mathbf{w}_h)|}{\|\mathbf{w}_h\|_{1,s}} \leq C\ h^{r'+2-d}\ |s|\ \left[\ |s|^2\ |\mathbf{v}_h|_{r-1,\Omega} + \sum_{m=0}^{d-1} h^m\ |||\mathbf{v}_h|||_{r+3+m-d,\Omega}\ \right].$$

with $r' = r$ *if* $Q = G$, $\qquad r' = r-1$ *if* $Q = GL$

Using the results of lemmas 2.9 and 2.5 in Strang's lemma, we obtain the main result of this section. For the sake of simplicity of our exposition, we shall restrict ourselves to the case of $r \geq 3$, which permits us to simplify the proof (we can use (2.36)). The same type of result with $r < 3$ still holds but one has to change the regularity requirements on $\mathbf{u}$, see for instance [15].

Theorem 2.10. *Let* $\mathbf{u}$ *be the solution of (2.5) and* $\mathbf{u}_h$ *the solution of (2.6). If the data* $\mathbf{f}$ *is smooth enough such that* $\mathbf{u} \in \mathbf{H}^{r+2}(\Omega)$ *and* $r \geq 3$*, then there exists* $\mathcal{C} > 0$ *such that*

$$\begin{aligned}\|\mathbf{u}-\mathbf{u}_h\|_{1,s} \leq\ & C\ h^r\ \left(1+\frac{|s|}{\eta}\right)\ \Big[\ |s||\mathbf{u}|_{r,\Omega} + |\mathbf{u}|_{r+1,\Omega}\Big] \\ & +\ C\ h^{r'+2-d}\ \ \frac{|s|}{\eta}\ \left(\ |s|^2\ |\mathbf{u}|_{r-1,\Omega} + \sum_{m=0}^{d-1} h^m\ |||\mathbf{u}|||_{r+3+m-d,\Omega}\ \right)\end{aligned}$$

with $r' = r$ *if* $Q = G$, $\quad r' = r-1$ *if* $Q = GL$

Proof. We apply the lemma 2.9 for $v_h = \pi_h \mathbf{u}$.

$$\sup_{\mathbf{w}_h \in \mathbf{U}_h^r} \frac{|(a-a_h)(s;\pi_h\mathbf{u},\mathbf{w}_h)|}{\|\mathbf{w}_h\|_{1,s}} \leq C\ h^{r'+2-d}\ |s|\ \left[\ |s|^2\ |\pi_h\mathbf{u}|_{r-1,K} + \sum_{m=0}^{d-1} h^m\ |||\pi_h\mathbf{u}|||_{r+3+m-d,K}\ \right] .$$

It is known that for $l \geq 2$, the interpolation operator π_h has the following continuity property in $H^l(\Omega)$:

$$|\pi_h\mathbf{u}|_{l,\Omega}\ \leq\ C|\mathbf{u}|_{l,\Omega} \tag{2.36}$$

Since $r \geq 3$, $r-1 \geq 2$ and $r+3+m-d \geq r+3-d \geq r \geq 2$, we get

$$\sup_{\mathbf{w}_h \in \mathbf{U}_h^r} \frac{|(a-a_h)(s;\pi_h\mathbf{u},\mathbf{w}_h)|}{\|\mathbf{w}_h\|_{1,s}} \leq C\ h^{r'+2-d}\ |s|\ \left[\ |s|^2\ |\mathbf{u}|_{r-1,\Omega} + \sum_{m=0}^{d-1} h^m\ |||\mathbf{u}|||_{r+3+m-d,\Omega}\ \right] . \tag{2.37}$$

Now, we can apply the Strang Lemma, with $v_h = \pi_h \mathbf{u}$, and combine the error of interpolation 2.14 and the error of quadrature 2.37, to obtain the theorem. ∎

B. Error estimates in the time domain

In this section we come back to the error analysis linked to the approximation of (1.2) by (1.26). For this, the idea is to use the properties of the inverse Laplace transform and the estimates of the previous section.

A time domain functional framework

We introduce the space of finite energy solutions in an interval $[0,T]$

$$V_{T,0} = \{v \in H^1(\ 0,T;L^2(\Omega)\ ) \cap L^2(\ 0,T;H^1(\Omega)\ ) \text{ such that } v(x,0) = 0 \quad \forall\ x \in \Omega\}\ ,$$

and more important, the corresponding finite energy norm $|||.|||_{1,T}$ on $V_{T,0}$:

$$|||v|||_{1,T}^2 = \int_0^T \left[\ |\frac{\partial v}{\partial t}|_{0,\Omega}^2 + |\nabla v|_{0,\Omega}^2\ \right]\ dt\ , \tag{2.38}$$

with which we shall measure the error.

We shall also use the semi-norms

$$\forall\ v \in H^l(O,T;H^m(\Omega)) \quad , \quad |v|_{m,l,T}^2 = \int_0^T |\frac{\partial^l v}{\partial t^l}|_{m,\Omega}^2\ dt\ . \tag{2.39}$$

Finally, the use of Laplace transform naturally leads to work with the following weighted Sobolev spaces, parameterized by $\eta > 0$, of functions of $x \in \Omega$ and $t \in \mathbb{R}^+$:

$$H_\eta^l(\mathbf{R}^+;H^m(\Omega)) = \Big\{ v \in L^2_{\rm loc}(\mathbf{R}^+;H^m(\Omega))\ /\ \frac{\partial^p v}{\partial t^p} e^{-\eta t} \in L^2(\mathbf{R}^+;H^m(\Omega))\ ,\ 0 \leq p \leq l \Big\}\ . \tag{2.40}$$

Finally, let us briefly recall some definitions related to the Laplace transform of a distribution. Let E be a Hilbert space and define

$$LT(\eta,E) = \{f \in \mathcal{D}'_+(E) \text{ such that } e^{-\eta t} f \in \mathcal{S}'_+(E)\}\ .$$

where $\mathcal{D}'_+(E)$ and $\mathcal{S}'_+(E)$ denote, as usual, the sets of distributions and tempered distributions on $\mathbf{R}$ with values in E and support in $[0,\infty[$ (see [8]). Then, for $s \in \mathbb{C}^+$, we denote by $\mathcal{F}_L$ the Laplace transform on $LT(\eta, E)$:

$$f(x,t) \xrightarrow{\mathcal{F}_L} \mathcal{F}_L(f(x,t)) = \mathbf{f}(x,s) = \frac{1}{2\pi}\int_0^{+\infty} f(x,t)e^{-st}\,dt\ .$$

The key tool for our analysis will be the Plancherel theorem:

$$\text{For } s \in \mathbb{C}^+ \text{ and } w \in LT(\eta, L^2(\mathbf{R})), \quad \int_{\mathbf{R}} \mathbf{w}^2(s)\,d\omega = \int_0^{+\infty} w^2(t)e^{-2\eta t}\,dt\ .$$

The time domain error estimates

Is is straightforward to transform theorem 2.10 into the following time domain error estimates (for simplicity we restrict ourselves to $r \geq 3$ but the result extends to $r \leq 2$ provided some minor changes in the regularity requirements for the solution).

Proposition 2.5. *Let $r \geq 3$, u and u_h be the solutions of (1.2) and (1.26). Assuming that u has the regularity :*

$$u \in \bigcap_{l=1}^{3} H^l_\eta(\mathbf{R}^+; H^{r+2-l}(\Omega)) \quad \cap \quad \bigcap_{m=0}^{d-1} H^1_\eta(\mathbf{R}^+; H^{r+2-m}(\Omega)). \tag{2.41}$$

One has the error estimate with $r' = r$ if $Q = G$ and $r-1$ if $Q = GL$:

$$\begin{aligned}
&\Big[\int_0^{+\infty}\Big(|\frac{\partial(u-u_h)}{\partial t}|^2_{0,\Omega} + |u-u_h|^2_{1,\Omega}\Big)e^{-2\eta t}\,dt\Big]^{1/2} \leq \\
&\leq C\,h^r\ \Big[\int_0^{+\infty}\Big(|\frac{\partial u}{\partial t}|^2_{r,\Omega} + |u|^2_{r+1,\Omega}\Big)e^{-2\eta t}\,dt\Big]^{1/2} \\
&+ \frac{C}{\eta}\,h^r\ \Big[\int_0^{+\infty}\Big(|\frac{\partial^2 u}{\partial t^2}|^2_{r,\Omega} + |\frac{\partial u}{\partial t}|^2_{r+1,\Omega}\Big)e^{-2\eta t}\,dt\Big]^{1/2} \\
&+ \frac{C}{\eta}\,h^{r'+2-d}\ \Big[\int_0^{+\infty}\Big(|\frac{\partial^3 u}{\partial t^3}|^2_{r-1,\Omega} + \sum_{m=0}^{d-1} h^m\,|||\frac{\partial u}{\partial t}|||^2_{r+3+m-d,\Omega}\Big)e^{-2\eta t}\,dt\Big]^{1/2}
\end{aligned} \tag{2.42}$$

Proof. The proof is very classical, so we just give the main idea. As usual, remarking that the term s^α in the time harmonic domain correspond to $\frac{\partial^\alpha}{\partial t^\alpha}$ in the time domain, we just have to apply the Plancherel theorem to the inequality in the Theorem 2.10 (for instance see [15] for rigourous arguments). ∎

We can use a classical technique presented in [15] to get error estimates on a finite time interval $[0,T]$. The arguments are based on causality property and to the optimal choice of the parameter η with respect to the time T. As the proof is a repetition of the one given in [15], we restrict ourselves to state the final result.

Theorem 2.11. *Let $r \geq 3$. Assuming that the data f is smooth enough such that $u \in H^{r+3}(\Omega \times [0,T])$. Then, if u_h is solution of (1.26), we have the following error estimate, with $r' = r$ if $Q = G$ and $r-1$ if $Q = GL$:*

$$
\begin{aligned}
&|||u - u_h|||_{1,T} \leq \\
&\leq C\, h^r \left[\int_0^T \left(|\frac{\partial u}{\partial t}|^2_{r,\Omega} + |u|^2_{r+1,\Omega} \right) dt \right]^{1/2} \\
&+ C\, T\, h^r \left[\int_0^T \left(|\frac{\partial^2 u}{\partial t^2}|^2_{r,\Omega} + |\frac{\partial u}{\partial t}|^2_{r+1,\Omega} \right) dt \right]^{1/2} \\
&+ C\, T\, h^{r'+2-d} \left[\int_0^T \left(|\frac{\partial^3 u}{\partial t^3}|^2_{r-1,\Omega} + \sum_{m=0}^{d-1} h^m \, |||\frac{\partial u}{\partial t}|||^2_{r+3+m-d,\Omega} \right) dt \right]^{1/2}
\end{aligned}
\tag{2.43}
$$

The main results are summarized in the table available in figure 1.

III. NUMERICAL RESULTS

In this section, we present some numerical results to confirm and illustrate the error estimates established in the previous section. More precisely, we provide some numerical tests of convergence to illustrate the quadrature error estimates (2.20) and (2.31). To this end, we considere the domain $\Omega = [0,1]^d$ and $\mathcal{T}_h$ the mesh associated with the step h. As the theoretical results are obvious and well known in the case of an uniform mesh (cf. remarks 2.2 and 2.3), we focus our experiments on the case of a non-uniform mesh. To this end, we construct a non-uniform mesh from a mesh made of triangles or tetrahedron, splitting each triangle or (tetrahedron) into 2 (or 3) quadrilaterals (or hexahedra). We illustrate this approach in figure 4. To illustrate theorems 2.31 and 2.6,

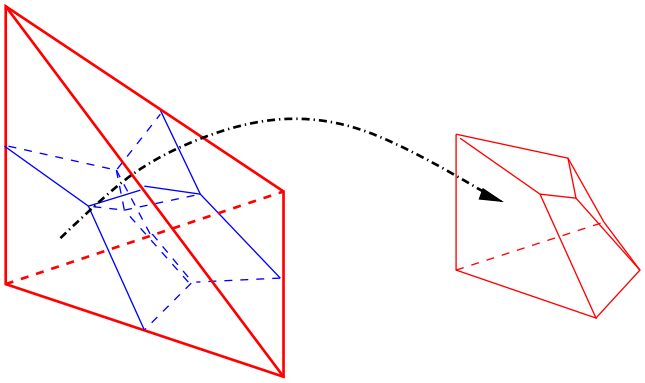


FIG. 4. How to cut tetrahedra in hexahedra.

we wish to compute successively the quantities

$$
E^m(h) = \sup_{w_h \in U_h^r} \frac{e_h^m(\, \Pi_h^r u \,,\, w_h \,)}{\|w_h\|} \quad \text{and} \quad E^s(h) = \sup_{w_h \in U_h^r} \frac{e_h^s(\, \Pi_h^r u \,,\, w_h \,)}{\|w_h\|} \,,
$$

where u is a very smooth fonction. To obtain reliable results, we do not compute these errors with only one smooth function but with a family of smooth functions. We proceed as follows: if we use $\mathcal{Q}^r$ finite elements, we choose an integer $R \geq r$ and we compute

$$E^m(h) = \max_{v_h \in \mathbb{M}_R} \sup_{w_h \in U_h^r} \frac{e_h^m(\ \Pi_h^r v_h\ ,\ w_h\ )}{\|w_h\|} \quad \text{and} \quad E^s(h) = \max_{v_h \in \mathbb{M}_R} \sup_{w_h \in U_h^r} \frac{e_h^s(\ \Pi_h^r v_h\ ,\ w_h\ )}{\|w_h\|}$$

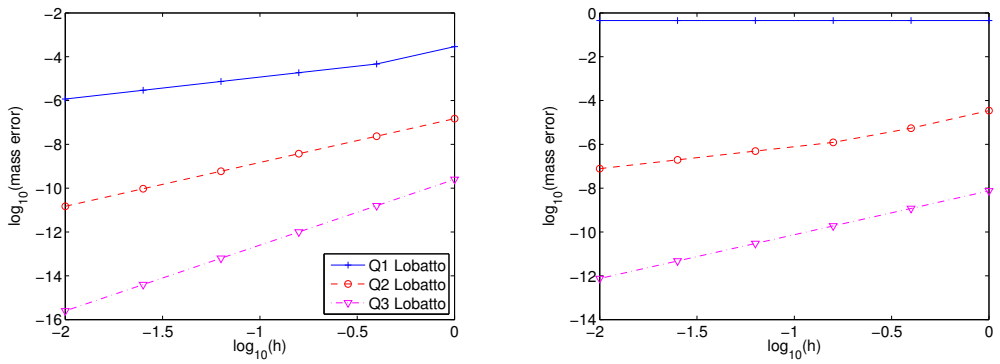
where $\mathbb{M}_R$ is the canonical basis of the polynomial space $\mathbb{Q}_R$. We observed that, in practice, it is sufficient to take $R = r$, because the worst order of convergence is observed for low-degree polynomials.

A. Concerning the error made on the mass term

We present in figures 6 the quantity $log(E^m(h))$ as a function of $log(h)$ from the order $r = 1$ to the order $r = 3$ and we sum up the rates of convergence in the figure 5 for the different cases (uniform or non-uniform mesh, 2D or 3D). One can see that the error

	2D.	3D
uniform mesh	$E^m(h) = O(h^{r+1})$	$E^m(h) = O(h^{r+1})$
non uniform mesh	$E^m(h) = O(h^r)$	$E^m(h) = O(h^{r-1})$

FIG. 5. Rates of convergence concerning the mass term for a Gauss-Lobatto quadrature rule



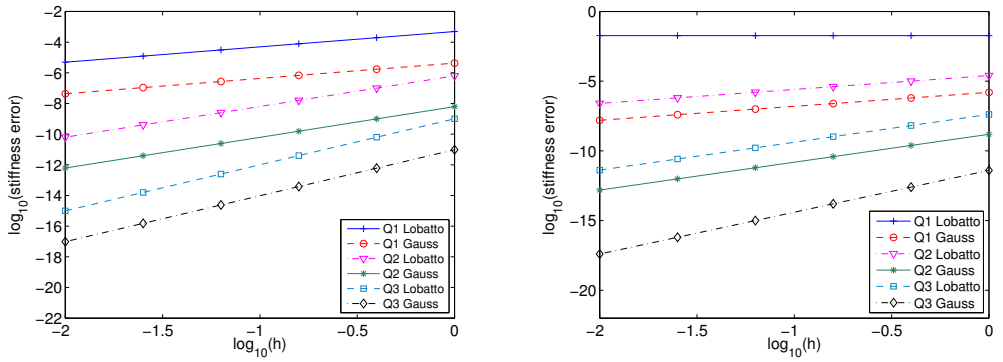
(a) The 2D case (b) The 3D case

FIG. 6. Convergence curves for the mass matrix using Gauss-Lobatto: non uniform meshes.

estimate (2.20) is optimal because it exactly agrees with the numerical results. Moreover, we can see in figure 6 that the order $r = 1$ is not consistent in 3D, as predicted by the theory.

B. Concerning the error made on the stiffness term

In the same way, we present on the figure 7 the quantity $log(E^s(h))$ as a function of $log(h)$ from the order $r = 1$ to the order $r = 3$, for both Gauss and Gauss-Lobatto quadrature rules.



(a) The 2D case (b) The 3D case

FIG. 7. Curves of convergence for a non uniform mesh, concerning the error of quadrature on the stiffness matrix by using Gauss-Lobatto or Gauss rules.

Computation with a Gauss quadrature rule

Figures 7 and 8 (the symbol ∞ corresponds to the case of exact integration) shows that if the numerical results are in agreement with the estimate (2.31) in the 2D case, this is no longer the case in three dimensions. Indeed, (2.31) provide the rate of convergence $O(h^{r-1})$ while the numerical experiments provide the rate $O(h^r)$. (cf. the figure 8).

	2D.	3D
uniform mesh	∞	∞
non uniform mesh	$E^m(h) = O(h^r)$	$E^m(h) = O(h^r)$

FIG. 8. Rates of convergence concerning the stiffness term for a Gauss quadrature rule

We also consider some numerical tests in the case where the stiffness matrix is computed with a Gauss-Lobatto quadrature rule. Let us recall that, in practice, a such choice is often used because it decreases about by twice the complexity for a matrix-vector product with the stiffness matrix (see [10] for a detailed analysis). The table given in figure 9 provides the observed rates of convergence in this case that appear to be better than the ones predicted by the theory (except of course the estimate (2.35) for regular meshes which is optimal).

	2D.	3D
uniform mesh	$E^m(h) = O(h^r)$	$E^m(h) = O(h^r)$
non uniform mesh	$E^m(h) = O(h^r)$	$E^m(h) = O(h^{r-1})$

FIG. 9. Rates of convergence for the H^1 norm for a Gauss-Lobatto quadrature rule

C. Numerical simulations for the Helmholtz equation

To confirm the influence of the numerical integration on a physical case, we propose some numerical results on the following model problem ($\Omega = [0,1]^d$)

$$\left| \begin{array}{ll} -\Delta \mathbf{u} - s^2 \mathbf{u} = 0 & , \text{ in } \Omega \ , \\ \dfrac{\partial \mathbf{u}}{\partial n} - i\ s\ \mathbf{u} = \dfrac{\partial \mathbf{u}^{inc}}{\partial n} - i\ s\ \mathbf{u}^{inc} & , \text{ on } \partial\Omega \end{array} \right.$$

where $\mathbf{u}^{inc}(x) = e^{i\ s\ x}$ and $s = \pi$. We compute an approximate solution with $\mathcal{Q}^r$ finite elements and different choices for the quadrature rules.

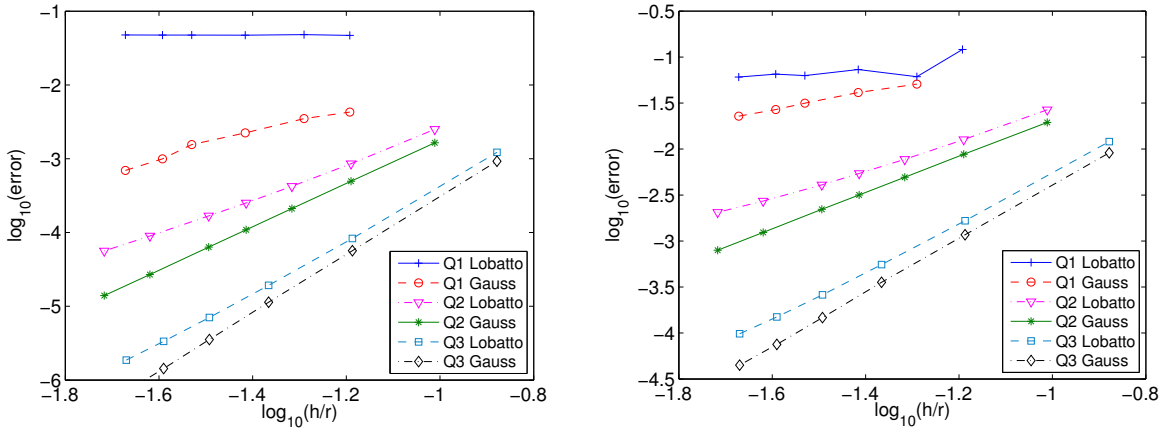


FIG. 10. At left, L^2 norm of the error $u - u_h$, at right error in H^1 norm. Gauss or Gauss-Lobatto quadrature are used, 3-D case. The model case is a cube meshed in tetrahedra split in hexahedra

Figures 11 and 10 illustrate the behaviour of the scheme if we use Gauss-Lobatto rules to integrate both matrices, or if we use the Gauss rules for both of them (in which case mass lumping is lost). One can see that we get a convergence in $O(h^r)$ in 2-D. The gain of accuray obtained by using Gauss rules is quite small, especially for $\mathbb{Q}_3$. For the 3-D case, we can see that we have a convergence in $O(h^{r-1})$ by using Gauss-Lobatto rules, while we get a convergence in $O(h^r)$ with Gauss rules. For $\mathbb{Q}_3$, the gain of accuray seems to be small.

Table in figure 12 gives the numerical estimations of the order of convergence when the mass matrix is computed with Gauss-Lobatto, and the stiffness matrix is computed with Gauss quadrature. Table in figure 13 gives the numerical estimation of the order of convergence when both matrices are integrated thanks to Gauss-Lobatto rules, for the 3-D unstructured case. It seems that there is no loss of order of accuracy between the two cases. One can see, that by using Gauss-Lobatto rule for mass and Gauss for stiffness, the loss of order of accuracy predicted by the theory can be seen with very fine meshes.

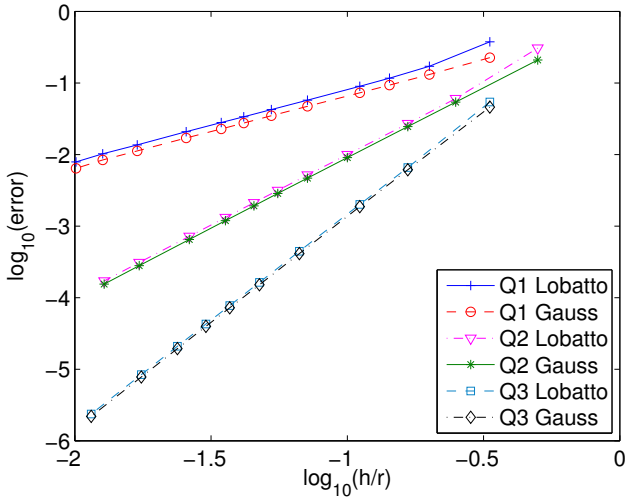


FIG. 11. Error in H^1 norm, by using Gauss or Gauss-Lobatto rules for the 2-D case. The model case is a square meshed in triangles split in quadrilaterals.

h	Error	Order
0.0642	0.174	-
0.0384	0.13	0.57
0.0256	0.152	-0.39
0.0213	0.157	-0.15

h	Error	Order
0.129	1.04e-2	-
0.0642	3.81e-3	2.58
0.0383	1.54e-3	1.76
0.0295	1.01e-3	1.58

h	Error	Order
0.195	1.4e-3	-
0.0966	1.79e-4	2.92
0.0771	9.34e-5	2.89
0.048	2.45e-5	2.82

FIG. 12. Estimate of order of convergence for the H^1 norm for the 3-D unstructured case. From left to right, $\mathbb{Q}_1, \mathbb{Q}_2 and \mathbb{Q}_3$. Integration of the mass matrix with Gauss-Lobatto rules and the stiffness with Gauss rules.

h	Error	Order
0.0642	0.121	-
0.0384	0.0733	0.98
0.0295	0.0628	0.58
0.0213	0.0607	0.107

h	Error	Order
0.129	1.27e-2	-
0.0642	4.07e-3	1.62
0.0383	2.05e-3	1.34
0.0295	1.49e-3	1.20

h	Error	Order
0.195	1.66e-3	-
0.0966	2.59e-4	2.64
0.0771	9.8e-5	2.38
0.048	5.05e-5	2.29

FIG. 13. Estimate of order of convergence for the H^1 norm for the 3-D unstructured case. From left to right, $\mathbb{Q}_1, \mathbb{Q}_2$ and $\mathbb{Q}_3$. Integration of the mass matrix and stiffness matrix with Gauss-Lobatto rules.

REFERENCES

1. Douglas N. Arnold, Daniele Boffi, and Richard S. Falk. *Approximation by quadrilateral finite elements.* Math. Comp., 71(239):909–922 (electronic), 2002.
2. Eliane Bécache, Abdelaâziz Ezziani, and Patrick Joly. *A mixed finite element approach for viscoelastic wave propagation.* Comput. Geosci., 8(3):255–299, 2004.
3. P. G. Ciarlet and P.-A. Raviart. *The combined effect of curved boundaries and numerical integration in isoparametric finite element methods.* In The mathematical foundations of the finite element method with applications to partial differential equations (Proc. Sympos., Univ. Maryland, Baltimore, Md., 1972), pages 409–474. Academic Press, New York, 1972.

4. Philippe G. Ciarlet. *The finite element method for elliptic problems*. North-Holland Publishing Co., Amsterdam, 1978. Studies in Mathematics and its Applications, Vol. 4.
5. Gary Cohen and Sandrine Fauqueux. *Mixed spectral finite elements for the linear elasticity system in unbounded domains.* SIAM J. Sci. Comput., 26(3):864–884 (electronic), 2005.
6. Gary Cohen, Patrick Joly, and Nathalie Tordjman. *Higher-order finite elements with mass-lumping for the* $1D$ *wave equation.* Finite Elem. Anal. Des., 16(3-4):329–336, 1994. ICOSAHOM '92 (Montpellier, 1992).
7. Gary C. Cohen. *Higher-order numerical methods for transient wave equations*. Scientific Computation. Springer-Verlag, Berlin, 2002. With a foreword by R. Glowinski.
8. Robert Dautray and Jacques-Louis Lions. *Analyse mathématique et calcul numérique pour les sciences et les techniques. Vol. 7*. INSTN: Collection Enseignement. [INSTN: Teaching Collection]. Masson, Paris, 1988. Evolution: Fourier, Laplace, Reprint of the 1985 edition.
9. Todd Dupont. L^2*-estimates for Galerkin methods for second order hyperbolic equations.* SIAM J. Numer. Anal., 10:880–889, 1973.
10. M. Durufle. *Intégration numérique et éléments finis d'ordre élevé appliquées aux équations de Maxwell en régime harmonique.* PhD thesis, Université Paris IX-Dauphine, 2006.
11. Vivette Girault and Pierre-Arnaud Raviart. *Finite element methods for Navier-Stokes equations*, volume 5 of *Springer Series in Computational Mathematics*. Springer-Verlag, Berlin, 1986. Theory and algorithms.
12. P. Grob. *Méthodes numériques de couplage pour la vibroacoustique instationnaire: Eléments finis spectraux d'ordre élevé et potentiels retardés.* PhD thesis, Université Paris IX-Dauphine, 2006.
13. P. Grob and G. Cohen. *Mixed higher order spectral finite elements for reissner-mindlin equations. SIAM J. Sci. Comput.*, In press, 2006.
14. J.P. Hennart. *On the efficient use of the finite element method in static neutron diffusion calculations.* Comput. Meth. in Nuclear Engee., 1:pp 185–199, 1979.
15. Patrick Joly, Gary Cohen, and Nathalie Tordjman. *Higher order triangular finite elements with mass lumping for the wave equation.* In Mathematical and numerical aspects of wave propagation (Mandelieu-La Napoule, 1995), pages 270–279. SIAM, Philadelphia, PA, 1995.
16. G. Cohen, P. Joly, J.E. Roberts, N. Tordjman *Higher order triangular finite elements with mass lumping for the wave equation. SIAM J. Numer. Anal.*, Vol 38(6), pp 2047–2078, 2001
17. W. A. Mulder, Chin-Joe-Kong and M. Van Veldhuizen. *Higher-order triangular and tetrahedral finite elements with mass lumping for solving the wave equation.* J. Engrg. Math., 35(4):405–426, 1999.
18. S. Pernet., X. Ferrieres *hp- a priori error estimates for a non-dissipative spectral discontinuous Galerkin method to solve the Maxwell equations in the time domain* In press Mathematics of Computation, 2006.
19. A. H. Stroud. *Approximate calculation of multiple integrals*. Prentice-Hall Inc., Englewood Cliffs, N.J., 1971. Prentice-Hall Series in Automatic Computation.
20. N. Tordjman. *Eléments finis d'ordre élevé pour l'équation des ondes*. PhD thesis, Université Paris IX-Dauphine, 1995.
21. O. C. Zienkiewicz. *The finite element method in engineering science*. McGraw-Hill, London, 1971. The second, expanded and revised, edition of The finite element method in structural and continuum mechanics.